DOCUMENT RESUME

ED 454 907 JC 010 495

AUTHOR	Armstrong, William B.
TITLE	Explaining Student Course Outcomes by Analyzing Placement Test Scores, Student Background Data, and Instructor Effects.
PUB DATE	2001-00-00
NOTE	24p.
PUB TYPE	Reports - Evaluative (142)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Ability Identification; *Academic Aptitude; Community Colleges; *Educational Testing; Outcomes of Education; Postsecondary Education; *Student Placement; Two Year Colleges
IDENTIFIERS	*California; University of California San Diego

ABSTRACT

This study examined the predictive validity of placement test scores with course grade and retention in English and mathematics classes. An increase in the use of testing to determine student eligibility for community college courses has prompted debate and litigation over the equity, access, and legal implications of these practices. In California, this has resulted in state regulations requiring that community colleges provide predictive validity evidence of test-score based inferences and course prerequisites. In addition, companion measures which supplement placement test score must be used for placement. However, for both theoretical and technical reasons, the predictive validity coefficients between placement test scores and final grades or retention in a course generally demonstrate a weak relationship. This investigation produced a model to explain variance in course outcomes using test score, student background data, and instructor differences. The model produced suggested that student dispositional characteristics explained the highest proportion of variance in the dependent variables. Including the instructor in the model added significantly to the explanatory power, suggesting that grading variation makes accurate placement more problematic. This investigation underscored the importance of academic standards as something imposed on the student by the college and not determined by the entering abilities of students. Contains 28 references. (JA)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 454 907

Explaining Student Course Outcomes by Analyzing Placement Test Scores, Student Background Data, and Instructor Effects

William B. Armstrong, Ph.D.
Director, Student Research and Information
University of California, San Diego
wbarmstrong@ucsd.edu

Abstract

Growth in the use of testing to determine student eligibility for community college courses has prompted debate and litigation over the equity, access, and legal implications of these practices. In California, this has resulted in state regulations requiring that community colleges provide predictive validity evidence of test-score based inferences and course prerequisites. In addition, companion measures which supplement placement test score must be used for placement. However for both theoretical and technical reasons, the predictive validity coefficients between placement test scores and final grades or retention in a course generally demonstrate a weak relationship. This study examined the predictive validity of placement test scores with course grade and retention in English and mathematics classes. The investigation produced a model to explain variance in course outcomes using test score, student background data, and instructor differences. The model produced suggested that student dispositional characteristics explained the highest proportion of variance in the dependent variables. Including the instructor in the model added significantly to the explanatory power suggesting that grading variation makes accurate placement more problematic. This investigation underscored the importance of academic standards as something imposed on the student by the college, and not determined by the entering abilities of students.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

W. Armstrong

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

TC010495

BEST COPY AVAILABLE

The open-access philosophy of the community colleges practically ensures that the students served will differ in their experiences, levels of education, and socio-economic status. Open-access combined with the growing national emphasis on improving student outcomes is a major challenge facing community colleges. Entry level placement testing is one method used by a growing number of community colleges to improve student outcomes. In California, legislation passed in the late 1980's mandated the use of placement testing as part of a comprehensive reform of college matriculation practices. The increasing use of placement testing to group students by ability led to legal challenges from groups concerned with testing practices and safeguarding open-access. In response, the State Chancellor's Office of the California Community Colleges (SCOCCC) implemented a series of regulations to govern the use of placement testing. Colleges were required to gather and report evidence of predictive validity of tests used to group and place students. In essence, community colleges must demonstrate that using tests to sort and place students enhances their likelihood for success in a course.

Predicting success in community college courses is made difficult by the differing characteristics of students. Sorting students by using cutting scores on a test may mask important individual characteristics and situations of students. In addition, poor predictions of student success or retention may occur because instructors -- acting as individual raters -- differ in their evaluation and retention practices. Prediction of student performance or retention in courses is thus often confounded when the criteria for success (such as course grade or retention) are unstable or influenced by other variables such as student and instructor characteristics.

Background

The Matriculation Act, passed in 1988 by the California legislature (California Code of Regulations (1990), promoted the use of placement tests in community colleges. Student opportunity and access were to be safeguarded by regulations that community colleges empirically demonstrate the reliability and validity of their placement tests and practices.

Growth in the prescriptive use of test scores to allocate access to community college courses and programs led to increasing anxiety by various social and advocacy groups concerned with the effects of testing. Thus the notion of test anxiety did not appear to apply just to students preparing to sit for an exam. A collective test anxiety could be found among state and local community college officials, faculty, testing experts, and activist groups regarding how standardized placement tests would be used, interpreted and applied in the community colleges. This collective test anxiety soon made its way into court.

In 1988, the Mexican-American Legal Defense Fund (MALDEF) filed a lawsuit alleging that several community colleges were using scores on placement tests to force students to take non-credit, remedial courses before they could enroll in courses carrying college credits that were transferable to four-year colleges and universities (Fields, 1988).

The MALDEF lawsuit was settled three years later. Partly as a result of the settlement, the State Chancellor's Office for the California Community Colleges (SCOCCC) agreed to require that every community college produce evidence of the criterion-related validity for each test they were using for placement (Cage, 1991). The evidence produced must conform to the Standards for Educational and Psychological

Testing (American Psychological Association, 1985). Community colleges must demonstrate that students have a demonstrably greater likelihood of success above a certain "cutting" score (minimum score on a test) than individuals falling below this score.

Despite the legal settlement and the implementation of admission, testing, and placement regulations over the last twelve years, the testing issue remains hotly debated. For example, nursing programs in the California community colleges are now required to admit all students who meet minimum requirements. Until this most recent ruling, nursing programs had relied on locally developed criteria for admitting students such as interviews, and ranking students on the basis of test scores and grades in prerequisite courses. Nursing programs must now empirically demonstrate that students falling below a certain test score are highly unlikely to succeed. The use of interviews and other "non-evaluative techniques," are prohibited (McCoy, 2000).

SCOCCC regulations also included a state-prescribed correlation coefficient of at least .35 between test score and the criterion (e.g., course grade). State regulations also required colleges to use companion measures in addition to placement test score in the placement decision. These other measures might include student high school GPA, employment hours worked, educational goal, and other data that may affect student course performance or retention (SCOCCC, 1992). Although the value of a scale that integrated student biographical data with test scores had been recommended by measurement scholars for several years (Cronbach, 1990; Anastasi, 1968), and advocated by professional organizations such as the American Psychology Association (APA)

(1985; 1999), many community colleges were slow to identify or adopt companion measures for placing students (Alkin and Freeman, 1992).

Establishing the predictive validity of test-score based inferences about student ability is often problematic, particularly in community colleges with existing placement programs. The pre-sorting of students into instructional levels truncates the range of the student test score and biographical characteristics available for study. Thus researchers often don't know how students would have performed had they been allowed to enroll in the higher level course.

A truncated distribution often results in lowered correlation coefficients between predictor and criterion (e.g., test score and course grade). However, this investigation used data from a period in which placement was non-mandatory. That is, students not found to meet the minimum test score or prerequisite for an English or mathematics course were allowed to enroll after advisement into a lower level course. This reduced, but did not eliminate, the restriction of the range of student ability and background in the courses under study.

Measurement scholars have noted that predictive validity studies are sometimes muddled or inconclusive because of unreliable or inadequate criterion data (APA, 1999; Popham, 1990; Sticht, Armstrong, et. al (1987); Cronbach, 1971). The validity of a placement system that sorts and classifies thousands of students is often reliant on the individual judgment provided by instructors, raters, or other evaluators of the student's performance.

To control for variance in the criterion variables of final grade this investigation included the instructor as a source of variance in the explanatory models. These models

were tested for both full-time and adjunct instructors to determine if instructor employment status was related to instability in grading.

Theoretical Models

The framework for this investigation relied on measurement theory and a review of the testing, measurement, and evaluation literature. Measurement theory was used to frame a study that recognized the difficulty of predictive validity when the criterion variable is a measure based on individual judgment, when student characteristics differ, and when the characteristics of the instructor differ. Two measurement theories; Point-to-Point Theory, and Aptitude-Treatment-Interaction (ATI) Theory guided the development of the course outcome model.

Predictive validity is essentially an attempt to approximate the future in the present. This is a central tenet of Point-to-Point Theory as described by Asher and Sciarrino (1974). Point-to-Point theory posits that predictive validity is enhanced to the extent that the correspondence between the skills measured on a test and the skills needed for success in a course are congruent. According to Asher and Sciarrino, "Information with the highest validity seems to have a point-to-point correspondence with the criterion" (p. 519). Point-to-Point Theory was found to be the most instructive in the formulation of research hypotheses and positing relationships between the independent variables of student test scores, demographic, dispositional, and situational characteristics and the dependent variables of course grade and retention.

ATI provided a theoretical framework for analyzing instructor effects on the criterion variables. ATI focuses on the processes and outcomes associated with student

referral to alternate educational treatments. This is analogous to the present situation in community colleges where students are referred to different levels in the curriculum. ATI focuses on the interaction of student traits, characteristics, and abilities, with particular forms of instructional treatments. This study extended the tenets of ATI by testing if instructional treatments, approaches, and grading standards within curricular levels vary by instructor.

Student Data

Constructs of student characteristics were identified and grouped using a model described by Cross (1981). Cross identified three primary factors related to course outcomes among adult students that were applicable to this investigation. These factors were situational, dispositional, and institutional variables. Situational variables included data pertaining to the present circumstances of the student such as employment hours, support for attending school, income, financial aid, part- or full-time attendance, and family responsibilities. Dispositional factors included cognitive, behavioral, and affective traits of the student such as self-efficacy, past experiences or performance in school, involvement in school activities, high school grade point average, high school preparation, and the perceived importance of attending school to the student. Institutional variables were also used by Cross (1981) in her analysis, but were not found to be useful to this investigation because all of the community colleges included in the study had similar policies and assessment practices. Thus there was little institutional variation to measure.

Literature Review

The literature review conducted for this investigation revealed that predictive validity studies conducted in community colleges and other post-secondary settings have yielded low correlation coefficients between placement test score and final grade (Hills, 1971; Cohen, 1987; Gabe, 1989; Hargis, 1990; Hughes and Nelson, 1991; Isonio, 1991; 1992; Rasor and Barr, 1993; College of the Canyons, 1994). However, the studies reviewed did not include student background information as part of the validation model.

With respect to instability in the criterion variable of final grade, there has been comparatively little evidence gathered that indicates how much error in the predictive validity of a test or other measure was contributed by individual judgments made by instructors, employers, supervisors, raters, or trainers. In discussing the usefulness of grades or marks as criterion data for evaluating placement decisions using tests, Cronbach (1971), made the following observations: "If teachers use different bases for judgment and some are more generous than others, throwing grades from several algebra teachers into a single distribution merely piles one source of error upon another" (p. 491). When a test fails to accurately predict student course performance, it is difficult to determine whether this observation derives from the test, the grading process, or other factors such as motivation, experience, and study habits, that also affect course performance or retention.

Research Questions

This investigation focused on three primary research questions:

1. Are placement tests highly predictive of course performance outcomes such as course grades?
2. How do student characteristics affect the prediction of course performance outcomes?

3. How do teacher characteristics affect the prediction of course performance outcomes?

Study Participants

To analyze the predictive validity of English and mathematics course placement, assessment data from all first-time students from three large-sized community colleges were gathered from the fall, 1995 semester. There were 3,925 students enrolled in one of the three levels of English courses (remedial or non-credit, Associate degree applicable, and transfer level), and 3,719 students enrolled in three levels of mathematics courses. The term was selected because of the temporary relaxation of placement rules that enabled some students to enroll in courses they would otherwise had not been eligible to enroll in.

Method

Testing research question one was done by deriving the predictive validity coefficients obtained through Pearson-Product-Moment correlations of test scores in reading, writing, and mathematics with the dependent variable of course grade. Course grade was recoded into a continuous variable with a score associated with each letter grade. Two computations of GPA were developed. One included GPA that excluded withdrawals (W's) from the analysis, while the second GPA measure counted W's the same as a failing grade. Statistical significance was set at the .05 level. This analysis was conducted for three curricular levels in English and mathematics.

The analysis conducted to answer research question two used student demographic, dispositional, and situational data in a regression model to determine the

amount of improvement that could be made in the prediction of student achievement and retention.

The course grade prediction model was constructed in a hierarchical approach. Test scores were first entered into the regression. Test scores were entered first because in practice, test scores were intended to predict course performance or serve as a proxy for entering ability. The subsequent ordering of the entry of variables into the model was intended to approximate the reality of placement practices in the colleges included in this study. This method was used to reflect both college matriculation policy and institutional practices in placement. The second block of variables entered into the equation consisted of demographic variables. Demographic variables were entered in this order because prior research suggests that certain demographic characteristics may mediate the predictive power of test scores. The third group of variables entered into the model was student dispositional data. The fourth block of variables entered into the model included the situational variables of the student.

The focus of research question three was the contribution of instructor differences in student evaluation to variance in course outcomes of final grade and retention. The analysis conducted to answer research question three controlled for test scores and student demographic, dispositional, and situational variation while noting the contribution of instructor grading practices to explaining or predicting variance in course performance. Anonymous instructor identification codes were entered into the model to detect the extent of error in the model associated with instructor variation.

The model explaining variance in final grade used the same sequencing of test score, demographic, dispositional, and situational variables used to provide evidence for

research question two. At the final stage of the regression model, the block of instructor data was entered into the regression model used to predict final grade. As each block of variables (test score, demographic, dispositional, situational, and instructor) were entered into the model, the change in the explanatory power of the of model was noted.

To determine the relation of faculty employment status (full- vs. part-time) to the outcome variables, two separate models were constructed. One model used only full-time faculty while the second model used only part-time faculty. The sequence of variable entry was the same for both models.

Findings

Research Question One

As was found in other predictive validity studies, although the correlation coefficients showed a statistically significant relationship between placement test scores and the dependent variables of course grade, the coefficients were too low to be of much practical significance. Within instructional levels (remedial or non-credit, Associate degree applicable, and transfer level) the coefficients varied with lower coefficients generally found at the lower levels of the curriculum. All correlation coefficients failed to meet the state-mandated .35 level for statistical validation. This was found for both of the dependent variables (GPA1 and GPA2) used to answer question one.

Research Question Two

Student test score, demographic, dispositional, and situational variables tended to contribute significantly to the models predicting course grade and retention. However, of all the variables entered, student dipositional variables tended to explain a greater amount

of variance in the dependent variables than did other student data, including test scores. Of particular value were data indicating the prior performance of the student in school such as high school GPA, grade in last English or mathematics course, and the number of years of English or mathematics taken in high school. This was true for both English and Mathematics courses. This analysis suggests that student dispositional data are more powerful predictors of student success than standardized test scores. This supported the theoretical construct posited by Point-to-Point theory. Past behavior is often the best predictor of future behavior.

Research Question Three

Including instructors in the predictive models significantly reduced error in the regressions. The amount of variance contributed by the instructor was generally 15-20% of the amount of variance in final grade. This suggested a relatively high degree of variation in grading practices by instructors. The instability in the criterion variable of final grade contributed to the low correlation coefficients found between test score and grade. Part-time instructors in English demonstrated a greater amount of grading variation than full-time instructors, however, no significant differences were found between full and part-time instructors in mathematics. Table 1 below summarizes the results for English courses for the three parts of the study and shows the relative contribution of each block of variables to the prediction of final course grade. Table 2 provides the same information for the mathematics courses analyzed.

Table 1: Factors Explaining Variance in English Course Grade by Full and Part-Time Instructor

Block of Variables	Multiple R	R squared	F	Significance of Change
Full-Time Instructors				
Placement Test Scores	.26	.065	21.2	p<.000
Demographic	.35	.122	6.9	p<.000
Dispositional	.43	.183	3.8	p<.000
Situational	.45	.206	3.4	p<.000
Instructor	.56	.312	2.7	p<.000
Part-Time Instructors				
Placement Test Scores	.24	.057	15.5	p<.000
Demographic	.39	.154	7.5	p<.000
Dispositional	.51	.264	5.0	p<.000
Situational	.52	.275	4.1	p<.000
Instructor	.68	.462	3.7	p<.000

The contribution of each block of variables as they entered the regression model can be noted in the column marked "R-square." This notes the proportion of variance accounted for by the entry of each particular group of variables. Of the student variables, although placement test scores are significantly related to final grade, they account for less than 7% of the variance in final grade. In contrast, dispositional variables contribute at least as much as test scores, and when used with demographic and situational variables, account for just over 20% of the variance in final grade. Entering the instructor in the model as an independent variable accounts for the greatest amount of variance in final grade (approximately 10%). The same pattern can be noted for part-time English instructors, however there is greater variance in grading practices among part-time faculty. The R-square increased by approximately 20% once the part-time instructors were entered into the regression model.

Table 2: Factors Explaining Variance in Mathematics Course Grade by Full and Part-Time Instructor

Block of Variables	Multiple R	R squared	F	Significance of Change
Full-Time Instructors				
Placement Test Scores	.14	.020	3.55	Not significant
Demographic	.47	.221	4.09	p<.000
Dispositional	.63	.397	2.71	p<.000
Situational	.65	.423	2.36	p<.000
Instructor	.77	.596	2.31	p<.000
Part-Time Instructors				
Placement Test Scores	.19	.037	9.66	p<.003
Demographic	.42	.177	4.73	p<.000
Dispositional	.54	.296	2.80	p<.000
Situational	.61	.374	3.07	p<.000
Instructor	.76	.577	2.78	p<.000

With respect to explaining variance in course grade in mathematics, a similar pattern emerges. Placement tests, however, were not found to be significantly related to course grades for full-time instructors. Student dispositional data accounted for almost 18% of the variance in final grade while placement test scores accounted for less than 5% of the variance for both full-and part-time instructors. Including the mathematics instructors in the regression model accounted for approximately 17-20% of the variance in final grade. This tends to confirm the presence of instability in the criterion variable of final grade. The finding of grading variation may help to explain the relatively low explanatory power of placement test scores.

Discussion

The predictive validity of test scores with respect to final grade is difficult to establish. Although California state regulations require colleges to demonstrate at least a .35 correlation, achieving this criterion for validity is difficult. This was found to be due to the differing characteristics and backgrounds of students and the grading practices used by individual instructors.

This investigation found that student dispositional and demographic variables had more explanatory power than did other variables, including test scores. This was found even though these data were self-reported by the student. Dispositional variables may be of greater predictive power final grade because they reflect the more enduring characteristics of the student that portend their likelihood for success. This finding is consistent with Point-to-Point theory, which formed the essential framework for this study.

This investigation also served to highlight the importance of the criterion for success. Misclassification of students as likely to fail is more than an incorrect inference derived from a test score. Misclassification may primarily be a function of who is assigning the grade. The interaction of student traits, instructional treatments, and instructor practices may have a greater effect on student performance than the skills measured by assessment tests. Poor prediction of performance or misclassification of students is thus exacerbated when the criterion for student success can vary depending on the class in which one enrolls. Thus a major finding of this study was that educational standards are maintained by the college, not determined by the entering ability of their students.

In the literature review conducted for this study, it was noted that thirty years ago, Sheldon (1970) wondered if accurate placement information about students could more easily be gleaned from a review of high school transcripts or even self-reported high school grade point averages than from mass administrations of standardized placement tests. Sheldon's supposition that student dispositional information is of more predictive value than standardized placement tests in community colleges tended to be confirmed in this study. Colleges seeking to improve their placement practices and increase the likelihood of accurate placement that improves student outcomes might want to consider using student dispositional data in the placement decision rather than strict reliance on placement test scores.

Based on the literature, theory, and findings from this study, practitioners and analysts intending to conduct predictive validity studies should expect to find low correlation coefficients. This may be due to several factors. Evaluating an existing assessment system where students are pre-sorted by ability group decreases the range of student abilities and lowers the correlation coefficients. Colleges wishing to conduct such studies may want to suspend mandatory placement for a semester or two, or lower the cutting score for entry into a particular course by five points or so. This will allow a wider range of student abilities into the classroom and assist with the conduct of the validity study.

To the extent possible, colleges should attempt to align the skills measured on placement tests with the behavioral objectives expected of students in the course. This was noted by Cohen and Brawer (1987):

Some of the more astute institutional researchers and faculty members have recommended using assessment tests that more closely approximate the behavior demanded of the students in class. They also seek reasonably common criteria for grading. Thus a testing program would have instructors closely involved in testing for the specific skills needed to succeed in a given class as identified by analyzing lectures, instructional materials, and assignments. Such a process has not been popularly adopted (p. 109).

Test content is of critical importance. State regulations that mandate a minimum correlation coefficient to establish the validity of a test do little to change the repeatedly observed findings that there are few common elements between placement test scores, and the factors that comprise final grade in a course. Implementing a placement test with only scant attention to the objectives and skills demanded of the course, greatly increases the likelihood that students, many of whom are capable of passing, will be misclassified as probable failures (Kangas, 2000). In other sectors such as the military or industry, institutions have developed tests that attempt to approximate the future in the present. For example, job sample tests use simulated exercises and situations that an applicant may confront in a training course or job. If college English courses emphasize writing, then one of the more obvious methods to approximate this objective is to have students complete a writing sample that could be scored by at least two English faculty to ensure reliability of placement. This might function as both a placement and diagnostic exercise, and provide greater information about the student than a norm-referenced score on a placement test. This may also help to promote faculty development for both full-time and adjunct faculty through review of the standards for various courses. Greater consistency in the application of academic standards may result through such shared activities as grading a writing sample.

The philosophy and practice behind placement programs typically assume that the instructional treatments are equivalent, and the criteria for success are reliable within treatments (courses). However, this study suggests that these assumptions are incorrect. Variance does not exist solely among the students. Variance was found to exist within instructional treatments, and among instructors, whether full time or adjunct. Prior to conducting local validation studies, practitioners may want to inspect the grades issued in a given semester by instructors teaching the same course section. If great variability is found (perhaps confirmed through analysis of variance procedures), steps should be taken to decrease this variance prior to conducting validation studies.

Short of comprehensive exit exams that mirror the skills assessed on placement tests, there will always be a relatively low point-to-point correspondence between test scores and final grade. The question then becomes, how much grading variation is tolerable? There are no references to give us the answer. Each college or department will have to discuss the issue. However, one guideline that emerges from the literature on inter-rater reliability is that the range of instructor GPA's should not exceed one-half of a grade. If this is found, it will be useful to share these findings with the academic department and begin to develop ways to achieve greater consistency. If grading variation is reduced and grades become more consistent through the enforcement of standards, higher predictive validity should result.

Do not place students using a single measure. When possible, gather dispositional data about the student that indicate their prior performance in school such as GPA, last English or math grade, commitment to school, and number of years of English or mathematics studied in high school. If the student has already taken at least 12 units of

community college courses, the college GPA may be a better predictor than self reported high school GPA. It is both verifiable and recent, and provides a better indicator of how a student will perform in class. However, even self-reported dispositional information has good predictive value. For example, Cohen and Brawer (1987) found that student self-assessment of ability in comparison with other students at their college was highly related to the content knowledge measured on the General Academic Assessment test. As noted by Cohen and Brawer, "Students know what they know." (p. 123).

The dispositional variables that best predict course success are not measured well nor generally gathered by standardized tests. If they are gathered, institutions often are not able to integrate the biographical information about the student in the placement decision for students. Colleges may want to consider methods to integrate dispositional data into the placement decision. This could be done through developing a composite scale, weighting the dispositional data and adding it to the test score, or using the dispositional data to direct a student to a higher or lower level course if their score is close to the cutting score.

This study sheds more light on the often repeated, but seldom understood, notion of academic standards (Astin, 1991). Many in higher education equate academic standards with the selectivity of the institution or program; that is, the caliber of the entering students is the indicator of the academic standards of the institution. This study suggests that academic standards might best be thought of as what the institution or instructor imposes on the students at the exit point, not the skills students bring with them at entry. Academic standards are brought to bear on the evaluation of student performance and are institutionally derived and defined (Cohen & Brawer, 1987). As

more students continue to enroll in the community colleges from increasingly diverse backgrounds and heritages, the imposition of consistent academic standards is the duty of the academic and administrative leaders of the college.

References

Alkin, M.C. and M. Freeman, (1992). Matriculation evaluation: Summary report 1991-1992. Northridge, CA: Educational Evaluational Associates.

American Psychological Association, (1999). Standards for educational and psychological testing. Washington, DC: Author.

American Psychological Association, (1985). Standards for educational and psychological testing. Washington, DC: Author.

Anastasi, A. (1968). Psychological testing (3rd edition.) New York: Macmillan.

Armstrong, W.B. (1999). Explaining community college outcomes by analyzing student data and instructor effects. Unpublished doctoral dissertation. University of California, Los Angeles.

Asher, J.J. and J.A. Sciarrino, (1974). Realistic work sample tests: A review. Personnel Psychology. 27. 519-533.

Astin, A.W. (1991). Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education. New York: Macmillan.

Cage, M.C. (1991, June 12). California community colleges system agrees to change role of testing. The Chronicle of Higher Education, p. A21.

California Code of Regulations (1990). Title 5. Statutes of 1990. Sacramento, CA: California State Printing Office.

Cohen, A.M. and F. Brawer, (1987). The collegiate function of community colleges. San Francisco, CA: Jossey-Bass.

College of the Canyons, (1994). Predictive validity of the APS writing and reading tests and validating placement rules for the APS writing test. Santa Clarita, CA: College of the Canyons, Office of Institutional Development. (ERIC Document Reproduction Service No. ED 376 915)

Cronbach, J.J. (1990). Essentials of psychological testing (5th ed.). New York: Harper Collins Publishers

Cronbach, (1971). Test validation. In R.L. Thorndike (ed.). Educational Measurement. (2nd edition) Washington, DC: American Council on Education.

Cross, K.P. (1981). Adults as learners: Increasing participation and facilitating learning. San Francisco, CA: Jossey-Bass.

Fields, C. (1988, June 1). California lawsuit challenges 2-year college' practice that restricts low test scorers to remedial courses. The Chronicle of Higher Education, P. A1.

Gabe, L.C. (1989). Relating college-level course performance to ASSET placement scores. (Institutional Research Report Abstract RR89-22). Fort Lauderdale, FL: Broward Community College. ERIC Document Reproduction Service No. 309 823).

Hargis, C. H. (1990). Grades and grading practices: Obstacles to improving education and to helping at-risk students. Springfield, IL: Charles C. Thomas Publishing.

Hughes, R.E., and C.H. Nelson, (1991). Placement scores and placement practices: An empirical analysis. Community College Review, 19 (1), 42-48.

Isonio, S.(1991). Implementation and initial validation of the APS English writing test at Golden West College: Evidence for predictive validity. Huntington Beach, CA: Golden West College. (ERIC Document Reproduction Service No. ED 345 781).

Isonio, S. (1992). Implementation and initial validation of the MDTP tests at Golden West College. Huntington Beach, CA: Golden West College. (ERIC Document Reproduction Service No. ED 345 782).

Kangas, J.A. (2000, Winter). Stirring the pot of statewide testing. Research and Planning Group for California Community Colleges News. Aptos, CA: Cabrillo College. p. 3.

McCoy, M.R. (2000, January 11). Open admissions tests California nursing programs. Community College Times. Washington, DC: American Association of Community Colleges.

O'Connor, M.C. (1989). Aspects of differential performance by minorities on standardized tests: Linguistic and sociocultural factors. Appearing in B. Gifford (Ed.) Test Policy and Test Performance: Educational, Language, and Culture. Boston, MA: Kluwer Academic Publishing, PP. 129-181.

Popham, W.J. (1990). Modern educational measurement: A practitioner's perspective. 2nd edition, Englewood Cliffs, NJ: Prentice Hall.

Rasor, R.A. and J.Barr, (1993). Refinement in assessment validation: Technicalities of dealing with low correlations and instructor grading variation. Sacramento, CA: American River College.

Sheldon, M.S. (1970). Entrance and placement testing for the junior college. Junior College Research Review. Dec. 1970, 5, 1-4.

State Chancellor's Office of the California Community College (SCOCCC), (1992). Standards, policies, and procedures for the evaluation of assessment instruments used in the California community colleges. Sacramento, CA: Matriculation Assessment Work Group. California Community Colleges.

Sticht, T.G, W.B. Armstrong, D.T. Hickey & J.S. Caylor (1986). Cast-off youth: Policy and training methods from the military experience. Westport, CT: Praeger Publishing.